

# **For Reference**

---

**NOT TO BE TAKEN FROM THIS ROOM**

## For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS  
UNIVERSITATIS  
ALBERTAEensis





Digitized by the Internet Archive  
in 2018 with funding from  
University of Alberta Libraries

<https://archive.org/details/evaluationoffe00roma>







THE UNIVERSITY OF ALBERTA

AN EVALUATION OF THE EFFECTIVENESS OF THE FIRST GRADE  
READINESS TESTING PROGRAM AS USED IN  
WEST JASPER PLACE PUBLIC SCHOOLS

by

Alexander Romaniuk

A THESIS  
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE  
OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL ADMINISTRATION

EDMONTON, ALBERTA

April, 1964



## ABSTRACT

Social pressure for the admission of technically under-age children to first grade has led the West Jasper Place Public School System to adopt a pre-school testing program for first-grade candidates between the ages of five and one-half and six years. The program is intended to measure the maturation of skills considered necessary for first-grade success and demands a predictive test of high validity.

Recently the Metropolitan Readiness Test in its two components, Number Readiness and Reading Readiness, was adopted as the standard for selection of under-age first grade candidates.

The Metropolitan Readiness Test has now been in use in West Jasper Place Public Schools long enough that a check on its validity becomes both possible and desirable. This study is consequently an attempt to determine the validity of the under-age first grade candidate selection system by statistical methods.

Using a sample of 444 candidates out of 610 students to whom the Metropolitan Readiness Test was administered before the 1961-1962 first-grade school year and using Marion Monroe Reading Test scores, Edmonton Public School Board Arithmetic Test (Grade 1) scores, and Teacher Ratings (determined in the last month of the first-grade school term) as criteria of achievement, this study uses Probability Tables and Canonical Variate Analysis to determine the correlation between predicted and measured achievement.

Results indicate that the correlation between achievement



predicted by the Metropolitan Readiness Tests and actual achievement by the standards of the criteria is of the order of .595, and in terms of a simple pass-fail criterion, the existing readiness testing program in West Jasper Place Schools has erred in only 13.5 per cent of the cases. In fact only 38 of the 441 under-age candidates in the study population were wrongly rejected in that they would have passed had they been accepted.

It has to be concluded that the Metropolitan Readiness Test is adequately fulfilling its intended function as the selection instrument for under-age first-grade candidates in West Jasper Place Public Schools, and that it merits retention in that function.



#### ACKNOWLEDGMENT

The writer wishes to express his appreciation to the many people whose cooperation made possible the completion of this thesis.

Particular thanks are due to members of his committee, Drs. H. T. Sparby, S. M. Hunka, and W. D. Neal, who were patient and meticulous critics; who offered encouragement and advice; and were always available and unfailingly helpful. The writer also acknowledges his indebtedness to Mr. A. J. Cropley for assistance with the statistical analysis used in this study.

His thanks are also given to the West Jasper Place Public School Board and their administrative staff for permitting the survey to be carried out in the system's schools and for the cooperation in the survey. Acknowledgment is also given to the principals, vice-principals, and teachers in West Jasper Place for participation in this survey.

Finally, no thesis could have been produced without the opportunity to attend the University of Alberta, an opportunity which was generously offered to and gratefully accepted by the author.



## TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION . . . . .	1
II. STATEMENT OF THE PROBLEM . . . . .	7
III. REVIEW OF RELATED LITERATURE . . . . .	10
Predictive Value of Tests . . . . .	10
Value of Teacher Evaluation of First Grade Students . .	14
Grouping of First-Grade Students . . . . .	18
Chronological and Mental Age as Factors in First	
Grade Achievement . . . . .	19
Sex of First Grade Pupils as a Factor in Achievement .	23
IV. STATISTICAL METHOD APPLIED TO STUDY . . . . .	25
Probability Table Method . . . . .	25
Canonical Variate Analysis Method . . . . .	26
V. SUMMARY OF STATISTICAL RESULTS . . . . .	29
Calculation of the Canonical Correlation	
Coefficient for Predictor Criteria Used	
in the Study . . . . .	39
Miscellaneous Calculations . . . . .	43
Projection of Number of Candidates Incorrectly	
Rejected by Predictors . . . . .	43
VI. SUMMARY OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS . .	45
Restatement of the Problem . . . . .	45
Findings . . . . .	45



CHAPTER	PAGE
Conclusions . . . . .	48
Recommendations . . . . .	49
BIBLIOGRAPHY . . . . .	50



## LIST OF TABLES

TABLE	PAGE
I. Correlations Between Metropolitan Readiness Tests and Metropolitan Achievement Tests . . . . .	12
II. Number of Children at Each Mental Age Level and the Percentage Making Average or Better Progress in Reading . . . . .	21
III. Probability Table Showing Frequency Distribution of Marion Monroe and Metropolitan Reading Scores . . . .	30
IV. Probability Table Showing Frequency Distribution of Metropolitan Reading and Edmonton Public School Board Arithmetic Scores . . . . .	31
V. Probability Table Showing Frequency Distribution of Metropolitan Reading and Teacher Rating Scores . . .	32
VI. Probability Table Showing Frequency Distribution of Metropolitan Numerical and Marion Monroe Scores . . .	33
VII. Probability Table Showing Frequency Distribution of Metropolitan Numerical and Edmonton Public School Board Arithmetic Scores . . . . .	34
VIII. Probability Table Showing Frequency Distribution of Metropolitan Numerical and Teacher Rating Scores . .	35
IX. Probability Table Showing Frequency Distribution of Metropolitan Reading and Metropolitan Numerical Scores . . . . .	36
X. Pearson Product-Moment Correlation Coefficients Between Predictor and Criteria Scores . . . . .	37







## CHAPTER I

### INTRODUCTION

The problem of first-grade school admission has long been a contentious one. It has always been understood that the criteria for admission of first-grade aspirants, because of the inevitable conflict between parental and school attitudes toward the matter, had to be as objective as possible. For much of the history of our schools, this has meant that a highly objective criterion such as the chronological age of the school aspirant has been used as the admission standard. The refinement of intelligence testing has, in more recent times, led to the relaxation of chronological age limits to permit the use of mental age as an auxiliary measure of admissability. Further sophistication of testing has taken place, and it is now a widespread practice in many educational systems to combine a chronological age standard of some latitude with tests that attempt to measure the extent of the maturation of those skills considered most necessary for achievement in school.

The history of the first-grade admission program in the West Jasper Place Public Schools has been true to this general pattern of development. The program is now using the Metropolitan Readiness Test, a test of the maturation of the skills required for early school work. However, this program, which replaced a mental-age testing program, is of fairly recent duration. It therefore seems advantageous at this time to examine the effect of the program in order to learn how well it is carrying out its intended function. It is for this purpose that the



following study has been instituted.

### Background to the Study

The first-grade readiness testing program of the West Jasper Place Public Schools consists of the administration of the complete Metropolitan Readiness Test to all candidates for first-grade admission. Candidates must be of the chronological age of at least five years six months, the legal school entry age in West Jasper Place being six years. A score of more than forty overall on the Metropolitan Readiness Test qualifies an under-age candidate for admission to first-grade. All children of legal age are admitted regardless of their test scores. The first-grade population resulting from this process in the school year 1961-62, for which this study is conducted, was 444 pupils.

Sixty-four candidates rejected because they obtained Metropolitan Readiness Test scores of less than forty were retested shortly before commencement of the first-grade school term. As a result five were admitted to first grade. Retesting used the Metropolitan Readiness Test in a repetition of the original test procedure.

Performances of the five retest acceptances are not separately taken into account in this study.

Use of the Metropolitan Readiness Test as the standard for under-age acceptability is relatively new in West Jasper Place. The Metropolitan Readiness Test supplanted the Detroit Test of Mental Ability in the autumn of 1960, and has been used exclusively since that time.

The test was originally administered in the first week of the



school term. In 1961 the testing period was changed to the last week in June and a three-day orientation period was introduced in order to accustom candidates to the school atmosphere preparatory to testing.

Additional use of the Metropolitan Readiness Test results is made during the first-year school term. The West Jasper Place Public Schools use an ability and achievement grouping system in the first-grade, and incipient groupings are made on the basis of test scores.

The first school year culminates in a pupil achievement evaluation program which is standardized throughout the West Jasper Place schools. This program is administered in the final week of the school year and consists of the following:

- 1) The Marion Monroe Reading Test;
- 2) The Edmonton Public School Board Achievement Test in Arithmetic (Grade I); and
- 3) Teacher Evaluation.

The Marion Monroe Reading Test. These tests have been designed to measure the extent to which pupils have progressed in the various aspects of reading emphasized by The New Basic Reading Program (grades 1 - 3). Each test serves both as a final test of mastery of the skills developed at a given level and as a readiness test for entering the next level of reading.

The standardization of the Marion Monroe Tests is based on a population of approximately 120,000 first, second and third-grade children from many representative American states. However, no figures on the reliability and validity of the test are available.



The Edmonton Public School Board Achievement Test in Arithmetic (Grade 1). This is a test of number skills, prepared by a committee of Edmonton teachers and used for promotion purposes in the Edmonton Public School system. The test is available from the Superintendent's Office, Edmonton Public School Board, Edmonton, Alberta, Canada. While norms have not been established for this test on a statistical basis, experiential results have, in the view of the Superintendent of the Edmonton Public School system, established a high correlation between scores from this test and other achievement tests of established reliability. This test is also considered to have a high correlation with predicted achievement as determined by the Metropolitan Readiness Test.

Teacher Evaluation. At approximately the same time as the evaluation tests are conducted, the teachers rate the pupils on overall achievement on a point scale ranging from one to five. Students assessed in the top 10 per cent of the class were given a rating of five. Students ranking in the next lower 15 per cent were classified as high and rated four. The next lower 50 per cent of students constituted the average group, rated three. Lowest rated groups corresponded in percentage to the highest groups, the 15 per cent nearest to average being rated low or two on the point scale. The lowest 10 per cent of students rated very low, or one on the scale. That is:

- 1) Very High--Scores made by the top 10 per cent of the class group.
- 2) High--Scores made by the next lower 15 per cent of the class group (75th to 89th percentiles).
- 3) Average-- Scores made by the middle 50 per cent of the class group



(25th to 74th percentiles).

- 4) Low--Scores made by the next lower 15 per cent of the class group  
(10th to 24th percentiles).
- 5) Very Low--Scores made by the lowest 10 per cent of the class group.

The teachers are expected to use the same general criteria but it is not known to what extent they attempt to equate their classes with others of the first grade.

These teacher ratings are included in the Promotion Sheets used in the West Jasper Place Public Schools, along with other confidential information. Lack of differentiation among pupils in the third, or average classification according to this rating scale makes the validity of correlating teacher ratings with the highly differentiated results of the Metropolitan Readiness Test a comparison of doubtful effect. A preponderance of either low or high average scores does not reflect in the teacher rating classifications, although it would be clearly indicated in the Metropolitan Readiness test results.

Another defect in the use of the teacher ratings for the purposes of this study lies in the extent to which the teachers may be influenced by the results of the Marion Monroe Reading Test and the Edmonton Public School Board Achievement Test in Arithmetic (Grade 1), in situations where the tests are taken before the teachers rate their classes. As well, it is not known how exactly the teachers comply with the recommended percentile for each rating classification. For that matter, it is impossible for the teachers, rating their own pupils only, to equate their percentiles with those applicable to a group of the



size used in this study.

Scores obtained by each pupil on the Metropolitan Readiness Test, the Marion Monroe Reading Test and the Edmonton Public School Board Achievement Test in Arithmetic (Grade 1), together with Teacher Ratings, are compiled by the school principals and forwarded to the office of the Superintendent of the West Jasper Place Public School Board. A sample of the form used in this compilation is shown below.

#### EVALUATION OF THE FIRST YEAR READINESS TESTING PROGRAM

SCHOOL: \_\_\_\_\_

NAME	SEX m f 1 0	AGE IN MONTHS	METROPOLITAN SCORES			MONROE READING		E.P.S.B.	TEACHER
			R.	N.	TOTAL	SCORE	LEVEL	ARITHMETIC	RATINGS
Doe, John	1	71	57	11	74	57	L.A.	71	2
Doe, Ruth	0	70	50	8	64	60	M.A.	64	3

This data, with additional School Board data respecting the age and sex of each pupil, constitutes the data used in this study.



## CHAPTER II

### STATEMENT OF THE PROBLEM

It is the purpose of this study to indicate whether the Metropolitan Readiness Test appears to be fulfilling its intended function in the West Jasper Place Public Schools. This function is dual. It includes providing an objective basis for rejection of those under-age first-grade candidates who do not demonstrate readiness, as well as predicting the achievement of accepted candidates. This study will deal only with the latter function of the Test.

There is reason to suspect that the Metropolitan Readiness Test may not be fulfilling its role accurately, in that the reports of some teachers suggest that an excessive number of changes in intraclass groupings based originally on the Test results must be made during the school year.

This study therefore explores the validity of the Metropolitan Readiness Test by making the following statistical correlations between predicted and measured achievement:

<u>Predicted Achievement</u>	<u>Measured Achievement (criteria)</u>
M.R.T. scores obtained in	Marion Monroe Reading Scores, E.P.
Pre-School entrance testing	S.B. Arith. (Gr. 1) Scores, and Teacher Ratings, determined during the last month of the first grade school year.

It therefore becomes the hypotheses of this study that:



- a) The intercorrelation between the predictor and criterion measure employed in the West Jasper Place Public School District grade one testing program will be significantly different from zero.
- b) Some of the measures employed will correlate significantly with sex of the students.
- c) It will be possible to obtain improved efficiency of predictors by the use of canonical correlation methods.
- d) The test weights associated with the coefficient of canonical correlation will indicate the differential efficiency of the various test measures in predicting grade one achievement.
- e) On the basis of the canonical correlation coefficient obtained, it will be possible to determine the proportion of errors associated with selection and rejection of under-age testees, and the proportion will be lower than that resulting from chance.

An indirect check on the validity of the Metropolitan Readiness Test also is available in the form of a correlation between the chronological ages of the study population and achievement as measured by the Marion Monroe Reading Test, the Edmonton Public School Board Achievement Test in Arithmetic (Grade 1), and Teacher Ratings. This check has the further merit of indicating the advisability of the practice of admitting under-age first-grade candidates that score higher than forty on the Metropolitan Readiness Test.

A statistically significant correlation between chronological age and measured achievement would suggest that the Metropolitan Readiness Test is overrating the capacity of the accepted under-age school pupils.



This study therefore makes a statistical correlation between chronological age and measured achievement in the form of:

- 1) Marion Monroe Reading Test scores,
- 2) Edmonton Public School Board Achievement Test in Arithmetic (Grade 1) scores, and
- 3) Teacher Ratings.

It also is possible from the available data to check one possible variable that might affect the ability of the Metropolitan Readiness Test to predict first-grade achievement. This variable is the sex of the pupil. Some studies indicate a definite difference between the levels of the performances of the sexes in the first grade, although others tend to contradict this indication. The causes of the differences in performance are in dispute, but there can be no doubt that where such differences exist, the ability of a readiness testing program to predict achievement accurately for one or the other sex must suffer unless different regression equations are used for each sex. In an effort to determine whether sex as a factor must be taken into account this study therefore correlates the sexes of the study population with their achievement as measured by:

- 1) Marion Monroe Reading Test scores,
- 2) Edmonton Public School Board Achievement Test in Arithmetic (Grade 1) scores, and
- 3) Teacher Ratings.



## CHAPTER III

### REVIEW OF RELATED LITERATURE

There exists a wealth of data and opinion related to this study. Virtually everything that has been written on the subject of readiness since Pestalozzi has at least some bearing on the subject of first-grade evaluation.

For that reason this review must be primarily selective and analytical. To accomplish this it is expedient that significant information be grouped according to the categories developed in this study. For convenience, these categories will be designated.

1. Predictive Value of Tests (Emphasizing the Metropolitan Readiness Test);
2. Value of Teacher Evaluation of First-Grade Pupils;
3. Grouping of First-Grade Pupils;
4. Chronological and Mental Age as Factors in First-Grade Achievement;
5. Sex of First-Grade Pupils as a Factor in Achievement.

#### I. PREDICTIVE VALUE OF TESTS

##### (EMPHASIZING THE METROPOLITAN READINESS TEST)

Objective evaluations of prospective first-grade entrants developed from a number of needs, principal among which were the need to limit the drastic number of first-grade failures<sup>1</sup> and the need arising

---

<sup>1</sup>Mary M. Reed, "An Investigation of Practice for the Admission of Children and the Promotion of Children from First Grade," Doctor's Dissertation, New York Teachers' College, Columbia University, 1927.



out of permissive first-grade entrance regulations, to provide an impartial evaluation of under-age and borderline first-grade prospects. The first testable and objective first-grade entrance determinant that came into common use was mental age. More refined study of first-grade pupils led to better realization of the particular skills that are necessary for successful learning in the first-grade. This study also led to curriculum modification<sup>2</sup> in some areas in de facto recognition of the state of these skills in the average school population. But the most significant result of the analysis of skills needed for first-grade achievement has been the development of tests which attempt to measure these skills more exactly than do mental age tests. The Metropolitan Readiness Test is one test that has been devised for this purpose. Eventually the results of readiness tests led to expansions of their significance into the fields of curriculum<sup>3</sup> and ability grouping. The tests' main validity, however, still appears to lie in prediction.<sup>4</sup>

The correlations (of the Metropolitan Readiness Test with the Metropolitan Achievement Test) are of such magnitude as to indicate the value of using the Metropolitan Readiness Tests as an aid in predicting school success at the first-grade level.<sup>5</sup>

---

<sup>2</sup>Arthur I. Gates, "Readiness for Beginning Reading," Readiness for Learning, Hildreth, G. (comp.), Bulletin of the Association for Childhood Education: Washington, 1941.

<sup>3</sup>Paul Witty, "What is Basic in Readiness for Learning?" Readiness for Learning, Hildreth, G. (comp.), Bulletin of the Association for Childhood Education: Washington, 1941.

<sup>4</sup>M. Lucille Harrison, Reading Readiness, Cambridge, Mass: Houghton, Mifflin Co., 1939.

<sup>5</sup>Metropolitan Readiness Tests - Reliability, New York and Tarrytown: Harcourt, Brace and World Inc., 1949.



Although the Metropolitan Readiness Test is of fairly recent origin, (1948), there has been considerable study of its reliability in predicting achievement.<sup>6</sup> Some indications of this reliability are shown below:

TABLE I

**CORRELATIONS BETWEEN METROPOLITAN READINESS TESTS  
AND METROPOLITAN ACHIEVEMENT TESTS\***

Metropolitan Readiness Tests		Metropolitan Achievement Tests	
		Average Reading	Numbers
Test 1	Word Meaning	.409	.441
Test 2	Sentences	.418	.424
Test 3	Information	.341	.400
Test 4	Matching	.486	.541
Test 5	Numbers	.520	.634
Test 6	Copying	.341	.403
Tests 1-4	Reading Readiness	.475	.530
Tests 1-6	Total Readiness	.534	.616

\*Data from Lebanon County, Pennsylvania (Number of cases, 487)

A study involving 487 pupils in Lebanon County, Pennsylvania, showed correlations between Total Readiness as scored by Metropolitan

---

<sup>6</sup>Metropolitan Readiness Tests "Validity", New York and Tarrytown: Harcourt, Brace and World Inc., 1949.



Readiness Tests and Metropolitan Achievement Tests, reading achievement of .534 and number achievement of .616<sup>7</sup> giving substance to the above quotation. Traxler<sup>8</sup> cites a test involving thirty-eight first-grade pupils in which readiness and total achievement correlated to the extent of .786.

The latter test is in line with what Gates demands of readiness testing, that is:

. . . should give a correlation of from .75 to .90 with reading ability achieved during the second semester.<sup>9</sup>

Many authorities recommend a testing program that is not solely dependent on a readiness test such as the Metropolitan Readiness Test. Hildreth holds that it is sometimes

. . . a good plan to use a general readiness test . . . with a good intelligence test . . . to increase the accuracy of measurement in predicting pupil progress,<sup>10</sup>

although Gates feels that

Comparatively few tests and appraisals suffice to give good results.<sup>11</sup>

<sup>7</sup> Ibid., p. 30.

<sup>8</sup> A. E. Traxler, The Reliability and Predictive Value of the Metropolitan Readiness Tests, Educational Records Bulletin No. 47, New York: Educational Records Bureau, February, 1947.

<sup>9</sup> Arthur I. Gates, "Readiness for Beginning Reading." Readiness for Learning, Hildreth, G. (comp.), Bulletin of the Association for Childhood Education: Washington, 1941, p. 15.

<sup>10</sup> Gertrude Hildreth, Readiness for School Beginners, Yonkers-on-Hudson, New York: World Book Co., 1950, p. 80.

<sup>11</sup> Arthur I. Gates, "Readiness for Beginning Reading," Readiness for Learning, Hildreth, G. (comp.), Bulletin of the Association for Childhood Education: Washington, 1941, p. 14.



The almost universal opinion is that formal objective testing is an invaluable predictive aid, but there are dissenting voices such as Gates<sup>1</sup>

The teacher and her methods of teaching are the decisive factors affecting the degree of success attained by pupils in beginning to read.<sup>12</sup>

and there is considerable agreement with Hildreth<sup>13</sup> and Gates<sup>14</sup> when they maintain that present tests and testing methods are far from perfect.

## II. VALUE OF TEACHER EVALUATION OF FIRST GRADE PUPILS

The importance of teacher evaluation lies in the generally accepted fact that a multitude of the factors that contribute to first grade achievement are not satisfactorily testable,<sup>15</sup> either because suitable good tests are not available or because the school testing program will not make allowances for a comprehensive readiness testing system.

There is substantial agreement that the factors affecting readiness are much as enumerated by Broom:

Psychological Factors - mental age, emotional stability, memory for a sequence of ideas, perception of relationships.

<sup>12</sup> Arthur I. Gates, "The Necessary Mental Age for Beginning Reading," Elementary School Journal, 1937, pp. 497-508.

<sup>13</sup> Hildreth, op. cit., p. 71.

<sup>14</sup> Gates, op. cit., p. 15.

<sup>15</sup> Jean D. Dey, Theory and Practice Governing the Time of School Entrance, Monographs in Education No. 4, The Alberta Advisory Committee on Educational Research, Edmonton: University of Alberta, 1960, p. 29.



Physical Factors - sensory discrimination, physical condition.

Intellectual Factors - concepts, word meanings, experience, organization and classification of ideas, concept of reading, experience with literature, attitudes.

Miscellaneous Factors - chronological age, left to right progression, social security.<sup>16</sup>

Of these, present readiness tests measure only a few of the psychological and the bulk of the intellectual factors. Typically, as Lazar says:

Tests give a mental age and intelligence quotient; . . . a total score in a reading readiness test . . . measuring specific ability to understand concepts, follow oral directions, understand stories read, match similar words, discriminate sounds, and copy words and numbers.<sup>17</sup>

The Metropolitan Readiness Test attempts to rate the following factors:

- 1) Understanding or comprehension of words,
- 2) Comprehension of phrases and sentences,
- 3) Vocabulary,
- 4) Recognition of similarities,
- 5) Number knowledge (number vocabulary, counting, ordinal numbers, recognition of written numbers, writing numbers, interpreting number symbols, the meaning of number terms, the meaning of fractional parts, recognition of forms, telling time and the use of numbers in simple problems),
- 6) Visual perception and motor control such as that required in learning

---

<sup>16</sup>M. E. Broom, and others, Effective Reading Instruction, New York: McGraw-Hill Book Company Inc., 1951.

<sup>17</sup>May Lazar, Determining Readiness for Reading, Educational Research Bulletin No. 6, Board of Education of the City of New York, 1943.



to write.

The large discrepancy between the number of factors involved in readiness and those factors testable is bridged somewhat by supplementary tests of the character of the Vineland Social Maturity Scale, Winnetka Scale for Rating School Behaviour and Attitudes, Falk's Inventory of Child Attitudes and by such questionnaires as that devised by Hurlock,<sup>18</sup> consisting of forty items relating to children's backgrounds of experience and skills.

Attempts have been made to use a broad testing program on a large scale. Ammens and Goodlad<sup>19</sup> report on the screening of early admissions in Pittsburgh where rather than depending on a limited testing system and limited criteria, psychologists used a wide range of data to evaluate readiness. The teacher in most school systems is in much the same relative position as the Pittsburgh psychologists.

A certain emphasis on teacher evaluation is to be expected from authorities who are, after all, teachers themselves. But the extent of their unanimity in considering teacher evaluation of first grade readiness as paramount is highly impressive, even though their opinion seems to lack statistical support.

Hildreth:

The teacher is in the best possible position to observe and rate the beginners. . . . Ratings of readiness based on teacher observations add to the prediction of pupil progress from tests, because they include variables that are not accounted for fully in tests.<sup>20</sup>

---

<sup>18</sup>E. Hurlock, Child Growth and Development, New York: McGraw-Hill Book Company Inc., 1949, Chapter 20.

<sup>19</sup>N. P. Ammens and J. I. Goodlad, "When to Begin", Childhood Education, Assoc. for Childhood Education International, Washington, 1955.

<sup>20</sup>Hildreth, op. cit., p. 48.



Broom:

At appropriate intervals the grade one teacher should evaluate each child, using the findings obtained by close observation and study and the results yielded by appropriate tests in order to determine his readiness for beginning reading.<sup>21</sup>

Harrison:

There are some factors which influence reading readiness for which there still are no objective measures. These factors for the most part, can be observed by the teacher and rated subjectively with a fair degree of reliability.<sup>22</sup>

Lazar:

The teacher is the key person in the situation.<sup>23</sup>

Witty:

Tests occasionally are useful in association with appraisals of the child's physical and social maturity.<sup>24</sup>

It follows, then, that teacher evaluation may be a highly useful prognostic instrument, and that, combined with a suitable testing program, it should give better predictions of first grade achievement than would a testing program alone. But it also follows that, since teacher evaluation is experiential, it is not a useful instrument of prediction until the teacher's familiarity with his pupils gives it value.

---

<sup>21</sup>Broom, op. cit., p. 102.

<sup>22</sup>Harrison, op. cit., p. 60.

<sup>23</sup>Lazar, op. cit., p. 9.

<sup>24</sup>Witty, op. cit., p. 10.



### III. GROUPING OF FIRST-GRADE PUPILS

Hildreth says:

There is considerable diversity of opinion and of practice with regard to the best age for first grade entrance, the best methods of grouping beginners and criteria to use as the basis of grouping. Brief test data and early observations are insufficient as a basis for permanent groupings.<sup>25</sup>

She feels that initial groupings should be purely tentative and that flexible groupings allow children to be shifted about as their needs require.

There is, nonetheless, some question as to the desirability of an easy attitude toward grouping first grade pupils. Changes of teachers as the child shifts from group to group are considered undesirable,<sup>26</sup> both for the child involved and by the child's parents. Group changes are also demonstrably undesirable from the school administration point of view, since they may contribute to an atmosphere of uncertainty within the school.

Is there a need for grouping? Grouping has, of course, been an essential part of school method since time immemorial. The grade system has been essentially a form of grouping on the basis of approximate chronological age and unassured achievement. It is now generally recognized that in the early years of schooling the pronounced differences in ability and achievement within the same approximate chronological age groups necessitate a process of sub-grouping.

---

<sup>25</sup>Hildreth, op. cit., p. 12.

<sup>26</sup>Lazar, op. cit., Chapter V.



Stability is a desirable aspect of grouping, and it is the aspect with which we are principally concerned in this study. The question of how best to obtain stability is answered by a number of authorities, who concur in the belief that measured readiness, or predicted achievement, is the most suitable standard for grouping.<sup>27,28</sup> Excessive instability of original groupings would therefore appear to be a direct result of unsatisfactory prognostication of pupil achievement occurring either as a result of an inferior readiness testing program or premature grouping, or a combination of both factors.

#### IV. CHRONOLOGICAL AND MENTAL AGE AS FACTORS IN FIRST GRADE ACHIEVEMENT

Dey found that although practices as to admission of first grade pupils on the basis of chronological age varied widely throughout the world,

Neither chronological age nor mental age are completely reliable as measures of the optimal<sup>29</sup> age at which all children can profit from formal instruction.

Bevington,<sup>30</sup> in a study of 640 pupils of the Edmonton Public System including a significant proportion of pupils under the age of six, could

---

<sup>27</sup>"Metropolitan Readiness Tests - Using the Test Results", New York and Tarrytown: Harcourt, Brace and World Inc., 1949.

<sup>28</sup>Hildreth, op. cit., p. 205.

<sup>29</sup>Dey, op. cit., p. 40.

<sup>30</sup>Bevington, W. "Effect of Age of Entrance into Grade One on Subsequent Achievement", Unpublished Master's Thesis, University of Alberta, 1957.



find no correlation between chronological age and achievement. Younger pupils achieved as much as older. However, it is often considered that a mental age of from six to six and one-half years under present methods of instruction, is essential for success in the first grade.<sup>31,32,33,34</sup> The studies of Hobson<sup>35</sup> and Gilmartin<sup>36</sup> on the success of under-age children tend to bear out this premise rather than refute it. Dean<sup>37</sup> provides a useful sample study of the correlation of mental age and achievement. His study was based on five first grade classes in Billings, Montana. Subjects were selected in order to provide as accurate a sampling as possible and variables regarding testing and teaching personnel were evaluated and considered not significant. Results are shown in Table II.

---

<sup>31</sup>Eleanor Hayes, "Why Pupils Fail", Educational Method, October, 1933, pp. 13-25.

<sup>32</sup>M. V. Morphett and C. Washburne, "When Should Children Begin to Read?" Elementary School Journal, 1931, pp. 496-503.

<sup>33</sup>W. D. Rice, "Some of the Educational Gains and Losses to Children Who Start to School at Different Ages," Master's Thesis, Chapel Hill: Duke University, 1934.

<sup>34</sup>Wilda Rosebrook, "Preventing Reading Deficiency", Elementary School Journal, 1936, pp. 276-280.

<sup>35</sup>James R. Hobson, "Mental Age as a Workable Criterion for School Admission," Elementary School Journal, 1948, pp. 312-321.

<sup>36</sup>Catherine E. Gilmartin, "Progress of Under-Age Children Admitted by Test to First Grade in Quincy Schools," Master's Thesis, Boston: Boston University, 1946.

<sup>37</sup>Charles D. Dean, "Predicting First Grade Achievement," Elementary School Journal, 1939, pp. 609-616.



TABLE II

NUMBER OF CHILDREN AT EACH MENTAL AGE LEVEL AND THE PERCENTAGE MAKING AVERAGE OR BETTER PROGRESS IN READING (Dean, 1939)

Mental Age Years and Months	Number of Children	Percentage Making Average or Better Progress
5.0 - 5.5	7	29
5.6 - 5.11	9	33
6.0 - 6.5	28	29
6.6 - 6.11	24	58
7.0 - 7.5	24	71
7.6 - 7.11	13	85
8.0 - 8.5	6	100

Correlation between Mental Age and Achievement = .62 (1939, p. 61)

It is notable that significant success was not obtained except in groups having a mental age of 6.6 or higher.

For the modern school system, the most significant aspect of the studies that have led to the decision that a mental age of as high as six and one-half years is desirable to ensure first grade achievement, is the inescapable conclusion that it might be desirable and even beneficial to postpone formal academic instruction until fully half of the first grade term has elapsed. It is not surprising that this coincides with other modern educators' theories regarding the delay of grouping programs and the use of the early portion of the first grade exclusively



as a readiness preparation period.

To supplement the argument that relatively advanced mental age is necessary for satisfactory first grade achievement, there is also evidence that maturation correlates highly with reading achievement.<sup>38,39,40</sup>

Opposing the conclusion that advanced mental age (and, by inference, advanced chronological age for large groups) is desirable for first grade achievement is Taylor's<sup>41</sup> study of Scottish School children. He found reading readiness and reading success can be affected by training and are not entirely or even predominantly dependent on maturation. Attempts to rationalize his results on the basis of criticism of his sample and other factors<sup>42</sup> are not entirely convincing. Gates<sup>43</sup> determined that where there was superior instruction, an abundance of suitable, individualized materials and relatively small classes, the

<sup>38</sup>Thomas H. Eames, "A Frequency Study of Physical Handicaps in Reading Disability and Unselected Groups," Journal of Educational Research, 1929, pp. 1-5.

<sup>39</sup>Willard C. Olsen, "Reading as a Function of the Total Growth of the Child", in William S. Gray (comp. and ed.) Reading and Pupil Development, Supplementary Educational Monographs No. 5, Chicago: University of Chicago Press, 1940.

<sup>40</sup>Elizabeth M. Fuller, "Peas in a Pod," Educational Leadership, 1953, pp. 302-307.

<sup>41</sup>Christian D. Taylor, The Effect of Training on Reading Readiness, Studies in Reading, Volume II, Scottish Council for Research in Education, London: University of London Press Ltd., 1948.

<sup>42</sup>Dey, op. cit., p. 69.

<sup>43</sup>Gates, op. cit., pp. 497-508.



supposed handicap of a low mental age, as low as five years, would not interfere with adequate reading achievement.

#### V. SEX OF FIRST GRADE PUPILS AS A FACTOR IN ACHIEVEMENT

The influence of sex on first grade achievement has always been regarded as difficult to measure because of the ancillary factors of bias in the predominantly female first grade teaching staffs and the supposedly feminine nature of overall first grade content. In fact the American Bureau of Education<sup>44</sup> found that in grade one girls did only slightly better than boys but the discrepancy in achievement increased in grades two and three. However, Samuels<sup>45</sup> examining one hundred pairs of boys and girls matched by mental ages, found achievement differences in favor of the girls great enough to have statistical significance.

Carrol<sup>46</sup> discovered statistically significant sex difference appearing during the reading readiness period, and Donnelly<sup>47</sup> found the greatest differences between the two sexes existed at the end of six months in

---

<sup>44</sup>"Pupils' Readiness for Reading Instruction upon Entrance to First Grade," City School Leaflet No. 23, U. S. Department of the Interior, Washington: Bureau of Education, December, 1926.

<sup>45</sup>F. L. Samuels, "Sex Differences in Reading Achievement," Journal of Educational Research, 36: 594-603.

<sup>46</sup>Marjorie W. Carroll, "Sex Differences in Reading Readiness", Master's Thesis, Boston: Boston University, 1941.

<sup>47</sup>Helen E. Donnelly, "A Study of Word Recognition Skills in Grade One," Master's Thesis, Boston: Boston University, 1940.



the first grade. Nila Smith concluded that:

Difference in physical, intellectual and emotional maturation is most frequently given as the reason why, on the whole, girls seem to be ready for successive levels of development in reading sooner than boys.<sup>48</sup>

---

<sup>48</sup> Nila Smith, "Readiness for Reading", Readiness for Reading and Related Language Arts, Bulletin, National Conference on Research in English, Elementary English, Chicago, 1950.



## CHAPTER IV

### STATISTICAL METHOD APPLIED TO STUDY

Two statistical methods, Probability Tables and Canonical Variate Analysis, were used to evaluate the study's results.

Probability tables establish relationships between variables. In the case of this study, the variables are designated predictors (i.e. Metropolitan Readiness Test reading scores and numbers scores) and criteria (i.e. Marion Monroe Reading Test Scores, Edmonton Public School Board Arithmetic Test scores and Teacher Ratings).

Canonical Variate Analysis determines the correlation between two composites, each made up of a number of tests - in this study, the same predictors and criteria subjected to probability table analysis.

#### Probability Table Method

Subjects' scores on each predictor were divided into class intervals. Criterion scores of subjects in each interval were then ascertained and recorded for each of the three criteria. Separate tables were then constructed for each predictor paired with each criterion, showing the frequency with which individuals who scored in each predictor interval fell into the class intervals of the criterion. Knowing the predictor score of an individual, it becomes possible to determine the likelihood of that individual's scoring in a given interval of the criteria.



### Canonical Variate Analysis Method

Cooley and Lohnes<sup>49</sup> have demonstrated the calculations and applicability of determining not simply the relationship between two single tests, but the relationship which can exist between one set of criteria and a second set of predictor tests. They have also shown that when the composite test is weighted so that the degree of relationship between such a battery and the single test is at a maximum, a useful correlation coefficient designated the multiple correlation coefficient is obtained. Extension of this procedure can also be made to yield a correlation between two optimally weighted composites, so as to yield the maximum degree of correlation between those composites. The resulting coefficient of correlation is referred to as a canonical correlation coefficient, and forms the basis of Canonical Variate Analysis.

A simple example of the use of canonical coefficient correlation can be illustrated when a number of tests of reading readiness are given to a group of school beginners, and the scores correlated with a number of reading scores at the end of a year's schooling. The readiness tests constitute the predictor battery, the final exam marks in reading the criteria.

The Pearson product-moment correlation coefficient, usually referred to as "r", in this case reflects the extent to which the subjects are placed in the same order by the criterion as the predictor

---

<sup>49</sup> W. W. Cooley and P. R. Lohnes, "Multivariate Procedures for the Behavioral Sciences," New York: Wiley, 1962.



indicated they would be.

In Multiple Correlation procedures, the aim is to find the correlation between a criterion (e.g. a final exam mark) and several predictors (e.g. an I.Q. test, plus a reading test, plus a number test, all administered earlier than the final exam). This approach yields the highest possible "r" between the single criterion and a compound score, obtained by multiplying each of the predictors by a calculated weight (referred to as a beta weight) and adding the multiplied predictors together. Beta weights are calculated as part of the same procedure which obtains the highest possible "r" between predictors and criterion, and must be employed if the highest possible "r" is to be obtained. Beta weights take account of how useful a particular test is in predicting. A high weight indicates that the test is a useful one in a predictive sense and measures a good deal of the criterion not measured by other predictors. If two tests measure the same thing, (i.e. suppose I.Q. and reading readiness are redundant measures) the better measure will be highly weighted, the poorer will receive a very low weight (even a negative weight, if the test is measuring something which is inversely related to success).

Canonical correlation is an extension of multiple correlation. In multiple correlation, the maximum "r" between a number of predictors combined with the best possible weightings is required; in canonical correlation, the highest possible "r" between the best possible combination of predictors and the best possible combination of a number of criteria is required (i.e. two sets of weights are needed, one to apply



to the predictors and one to the criteria). Criteria which measure nothing not better measured by other criteria, and useless predictors, receive low weightings.



## CHAPTER V

### SUMMARY OF STATISTICAL RESULTS

The data obtained in this study were used to construct probability tables according to the techniques explained earlier (page 25). This procedure yielded seven tables (Tables III - IX) showing the distribution of criteria scores for each predictor; using these tables it is possible to estimate the probabilities with which an individual scoring in a particular class interval or a particular predictor may be expected to score in the various class intervals of the various criteria. Tables III - V show the expected distribution of criteria scores for Metro. R Test, Tables VI - VIII show the relationship between criteria and Metro N scores, while Table IX shows the relative distribution of Metro R and Metro N tests.

Test scores were intercorrelated and a 't' test to test  $H_0: \rho=0$  against the alternative  $H_1: \rho \neq 0$  using  $\alpha=.01$  was calculated, using the formula

$$t = r \sqrt{\frac{N - 2}{1 - r^2}}$$

With 442 degrees of freedom, the critical value of "r" at the .01 level of confidence was found to be .111, while the critical value at the .05 level was .093.



TABLE III

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
 MARION MONROE AND METROPOLITAN  
 READING SCORES

				.006	.017	.041		
70-79								
60-69		.666	.375	.292	.416	.573	.895	
50-59			.375	.423	.355	.232	.062	
40-49			.125	.173	.167	.130	.020	
Monroe	30-39		.062	.076	.026	.028		
	20-29		.333		.038	.026	.017	
	10-19			.062				
	0-9							
Total		0	3	16	52	149	176	48
		0-9	10-19	20-29	30-39	40-49	50-59	60-69
								70-79

Metro. R.



TABLE IV

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
 METROPOLITAN READING AND EDMONTON PUBLIC  
 SCHOOL BOARD ARITHMETIC SCORES

			.019	.047	.102	.291	
90-99							
80-89		.333	.187	.096	.093	.340	.416
70-79			.250	.269	.228	.279	.208
60-69			.250	.292	.322	.176	.083
50-59		.333	.125	.134	.174	.084	
Arith.	40-49		.062	.096	.093	.017	
	30-39	.333		.076	.026		
	20-29		.062		.020		
	10-19			.019			
0-9		.062					
	Total	3	16	52	149	176	48
	0-9	10-19	20-29	30-39	40-49	50-59	60-69
	444						

Metro. R.



TABLE V

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION  
OF METROPOLITAN READING AND  
TEACHER RATING SCORES

			.019	.026	.119	.437	
5							
4		.666	.087	.096	.167	.301	
3			.312	.326	.369	.335	.354
Teacher							
2			.375	.442	.329	.221	.020
1		.333	.125	.115	.107	.022	
Total		3	16	52	149	176	48
							444
	0-9	10-19	20-29	30-39	40-49	50-59	60-69
						Metro. R.	



TABLE VI

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
 METROPOLITAN NUMERICAL AND  
 MARION MONROE SCORES

70-79		.006	.012	.044	
60-69	.277	.413	.550	.776	1.00
50-59	.293	.366	.297	.089	
40-49	.277	.146	.177	.074	
Monroe	30-39	.068	.046	.012	.014
	20-29	.068	.020	.018	
	10-19	.017			
	0-9				
Total		58	150	158	67
		0-5	6-10	11-15	16-20
					21-25

Metro. N.



TABLE VII

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
 METROPOLITAN NUMERICAL AND EDMONTON PUBLIC  
 SCHOOL BOARD ARITHMETIC SCORES

		.020	.075	.238	.818	
90-99						
80-89	.120	.120	.303	.417	.181	
70-79	.137	.286	.284	.223		
60-69	.241	.326	.196	.104		
50-59	.241	.140	.094	.014		
Arith.	40-49	.069	.093	.031		
	30-39	.103	.006	.012		
	20-29	.051	.006			
	10-19	.017				
	0-9	.017				
Total		58	150	158	67	11
		0-5	6-10	11-15	16-20	21-25

Metro. N.



TABLE VIII

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
 METROPOLITAN NUMERICAL AND  
 TEACHER RATING SCORES

5	.017	.026	.088	.328	.545
4	.086	.160	.284	.388	.454
3	.241	.366	.405	.179	
2	.431	.460	.196	.104	
1	.224	.080	.025		
Total	58	150	158	67	11
	0-5	6-10	11-15	16-20	21-25
				Metro. N.	



TABLE IX

PROBABILITY TABLE SHOWING FREQUENCY DISTRIBUTION OF  
METROPOLITAN READING AND METROPOLITAN  
NUMERICAL SCORES



The matrix of correlation between tests is shown in Table X. It is evident that all of these correlations are significantly different from zero at the .01 level.

TABLE X

PEARSON PRODUCT - MOMENT CORRELATION COEFFICIENTS  
BETWEEN PREDICTOR AND CRITERIA SCORES

	Test	1	2	3	4	5
1	Monroe	1.0	636*	716	312	389
2	Arithmetic		1.0	674	453	531
3	Teacher			1.0	426	502
4	Metro R				1.0	595
5	Metro N					1.0

\*All decimal points have been omitted in off-diagonal elements.

Furthermore, point biserial correlation coefficients were calculated between all tests and sex of the subjects. The point biserial coefficient, which is a form of Pearson product-moment correlation coefficient, expresses the degree of correlation between a continuous and a dichotomous variable. The point biserial correlation coefficients obtained are shown in Table XI.



TABLE XI

## POINT BISERIAL CORRELATIONS OF TESTS WITH SEX

Test	Metro R	Metro N	Monroe	Edmonton Arithmetic	Teacher Rating
r sex.test	-.077	-.026	-.109	.003	-.143

Of these correlations, only that between Teacher Ratings and Sex reached the .01 level of significance ( $r=.111$ ). Of the remainder only the correlation between Monroe Reading and Sex was significant at the .05 level ( $r=.093$ ).

Finally, product-moment correlations between age of the subjects and scores on each of the tests were calculated. These correlation coefficients are shown in Table XII.

TABLE XII

## PRODUCT-MOMENT CORRELATIONS OF TESTS WITH AGE

Test	Metro R	Metro N	Monroe	Edmonton Arithmetic	Teacher Rating
r age.test	.032	-.014	-.056	.006	-.067

None of these figures is significantly different from zero.



Calculation of the Canonical Correlation Coefficient for Predictors and Criteria Used in the Study

Calculation of the canonical correlation coefficient involves finding a set of weights such that, on the basis of those weights, it is possible to combine the predictors and criteria so that the reproduction of the criteria on the basis of the predictors is most accurately carried out. This involves finding solutions to the equation

$$Rx = x\lambda$$

where R is as shown below, x is the weight vector, with which is associated the scaling factor  $\lambda$ , whose magnitude reflects the degree to which it is possible to perfectly reconstruct the original criterion scores on the basis of the predictors, using a single weighting system contained within x. (i.e., it is intimately connected with the correlation between the two sets of scores, and is, in fact, equal to that correlation squared).

Calculations were carried out according to the technique suggested by Cooley and Lohnes.<sup>50</sup> To facilitate computations, the various tests were rearranged so that the left hand set of tests contained more tests than the right hand set (i.e. the criteria were written as tests 1, 2 and 3, the predictors as tests 4 and 5).

This rearrangement yielded the matrix of correlations shown in Table XIII.

---

<sup>50</sup>Cooley, op. cit., p. 36.



TABLE XIII

 PARTITIONED MATRIX OF INTER-CORRELATIONS  
 BETWEEN PREDICTORS AND CRITERIA

	Monroe	Edmonton Arithmetic	Rating	Metro R	Metro N	
1	1.00	.636	.716	.312	.389	
R <sub>11</sub>	2	.636	1.00	.674	.453	.531
	3	.716	.674	1.00	.426	.502
R <sub>21</sub>	4	.312	.453	.426	1.00	.595
	5	.389	.531	.502	.595	1.00
						R <sub>22</sub>

The correlation matrix was then partitioned as shown above, and R calculated on the following basis:

$$R = R_{22}^{-1} R_{21} R_{11}^{-1} R_{12}$$

R<sub>22</sub><sup>-1</sup> and R<sub>11</sub><sup>-1</sup> were calculated and found to be as shown below:

$$R_{22}^{-1} = \begin{bmatrix} 1.547 & -.921 \\ -.921 & 1.547 \end{bmatrix} \quad R_{11}^{-1} = \begin{bmatrix} 2.237 & -.627 & 1.176 \\ -.627 & 1.995 & -.897 \\ -1.176 & -.897 & 2.442 \end{bmatrix}$$

and hence;

$$R = \begin{bmatrix} .1107 & .1282 \\ .2070 & .2446 \end{bmatrix}$$

Now, since Rx = xλ, (R - λI)x = 0. For this to be true,

| R - λI | must be equal to 0.



$$\text{i.e. } \begin{vmatrix} .1107 - \lambda & .1282 \\ .2070 & .2446 - \lambda \end{vmatrix} = 0$$

$$\text{therefore, } (.1107 - \lambda)(.2446 - \lambda) - .2070 \times .1282 = 0$$

$$\text{solutions are } \lambda = .3536 \text{ or } .0017$$

The second root appeared so small relative to the first that the weights ( $x$ ) were calculated on the basis of  $\lambda_1 = .3536$ .

$$\text{i.e. } \begin{bmatrix} -.2429 & .1282 \\ .2070 & -.1090 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = 0$$

This matrix equation is satisfied by  $b_1 = .528$  and  $b_2 = 1.00$ .

The vector of weights  $[a_i]$  for the left-hand tests may be calculated by substituting into  $A_i = (R_{11}^{-1} \quad R_{12} \quad b_i) / \sqrt{\lambda_i}$

$$\text{This substitution yields } a_1 = \begin{bmatrix} -.167 \\ .903 \\ .728 \end{bmatrix}$$

Thus optimum relationship between the criterion battery and predictor battery can be obtained when the following weights are used:

$$\begin{array}{lll} .528P_1 + 1.00P_2 \text{ and } -.167C_1 + .903C_2 + .728C_3. \\ \text{Metro} \quad \text{Metro} \quad \text{Monroe} \quad \text{Edm.} \quad \text{Rating} \\ R \quad N \quad \quad \quad \quad \quad \text{Arith.} \end{array}$$

The canonical correlation associated with this weighting system is  $\sqrt{\lambda}$ ,

$$\text{i.e. } R_{\text{canonical}} = \sqrt{.3536} = .595$$

The significance of this correlation coefficient was tested as suggested by Cooley and Lohnes.<sup>51</sup> In order to use this test it is necessary to calculate  $\Lambda = \sum_{i=1}^q (1 - \lambda_i)$ .  $q < p$

<sup>51</sup>Cooley, op. cit., p. 37.



The value of  $\chi^2$  is then obtained from

$$\chi^2 = - \left[ N - .5(p + q - 1) \right] \log_e \Delta$$

where  $N$  = no. of subjects, and

$p$  and  $q$  = no. of predictors and criteria respectively.

Substituting into this equation,

$$\begin{aligned} \chi^2 &= - \left[ 444 - .5(6) \right] \log_e (1 - .3536)(1 - .0017) \\ \text{i.e. } \chi^2 &\approx 193. \end{aligned}$$

The number of degrees of freedom associated with this statistic is  $pq$ , i.e.  $df = 6$ . The critical value of  $\chi^2$  with 6 df ( $\alpha = .001$ ), is 22.46.

On the basis of this test it is reasonable to conclude that this root is significant beyond the .001 level of confidence. Similarly, a  $\chi^2$  test for the significance of the other root (.0017) was carried out.

In this case  $\chi^2 = 0.766$  (i.e. root is not significant).



### Miscellaneous Calculations

The multiple "R" between predictors and criteria was calculated, with unit weightings, and was found to be .551. This may be compared with the R canonical of .595 - some improvement of prediction is effected by the use of weightings, but it is only very slight. Furthermore, the canonical "R" with the Monroe Reading Test omitted is .584.

### Projection of Number of Candidates Incorrectly Rejected by Predictors

The number of children incorrectly placed on the basis of the predictors was calculated. Incorrect predictions may occur for two reasons - acceptance of children who are subsequently decelerated, and rejection of children who would have done satisfactory work if they had been accepted.

Based on the canonical correlation coefficient previously calculated and using the tables of Arbous<sup>52</sup> the proportion of correct and incorrect placements shown in Table XIV were obtained. These proportions are partly determined by the particular cut-off scores used in the West Jasper Place Public School District. Applying these proportions to the present sample, the results below were obtained.

Total no. of five-year-olds who were tested	- 441
---	-------

Selectees who were decelerated (i.e. errors in selection)	- 21
---	------

Selectees who passed (i.e. correct selections)	- 361
--	-------

Candidates who would have passed if they had been	
---	--

---

<sup>52</sup> A. G. Arbous, Tables for Aptitude Testers, Capetown: South African Council for Scientific and Industrial Research, 1952.



selected (i.e. they represent incorrect rejections)	- 38
Candidates who were rejected who would have	
failed anyway (i.e. correct rejections)	- 21

TABLE XIV

PROPORTIONS OF CHILDREN PLACED ACCORDING TO  
PREDICTION AND CRITERION TESTS

		CRITERION	
		Decelerate	Normal
P R E D I C T I O N	Select	B	A*
	Reject	.048	.819
		C	D
		.046	.087

---

\*Cells A and C represent correct decisions,  
Cells B and D represent errors in placement  
(i.e. 13.5 per cent)



## CHAPTER VI

### SUMMARY OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

#### Restatement of the Problem

Purpose of the study was to determine statistical correlations between predicted achievement and measured achievement in the case of 441 technically under-age first-grade candidates (aged five and one-half to six years), admitted to first grade of the West Jasper Place Public School System on the basis of scores obtained in a pre-school readiness testing program.

The Metropolitan Readiness Test in its two sections, Metropolitan Reading Readiness and Metropolitan Number Readiness, constituted the Predictor. The Marion Monroe Reading Test, Edmonton Public School Board Arithmetic Test (Grade 1) and Teacher Ratings constituted the criteria of measured achievement.

#### Findings

1. The intercorrelations between all predictor and criterion measures differed significantly from zero, beyond the .01 level of confidence (Table X). Consequently, Hypothesis A was supported by the data employed in this study.

2. The correlation between Teacher Ratings of students' performances and the sex of the students was significantly different than zero at the .01 level of confidence, while the correlation between Sex and Monroe Reading Test scores was significant at a lower level (.05). Neither of the other predictors correlated significantly with sex, nor



did the Edmonton Arithmetic Test scores. Hypothesis B was supported by the experimental data.

3. With unit weightings, the correlation between the predictor and criterion batteries was .551. However, this correlation was raised to .595 by combining the various measures according to the calculated canonical weights. Consequently, Hypothesis C was supported, although the improvement in prediction was slight.

4. The test weights calculated in the canonical correlation procedures were -.167 (Monroe Reading Test), .903 (Edmonton Arithmetic Test), .728 (Teacher Ratings), 1.00 (Metropolitan Number Test) and .528 (Metropolitan Reading Test) respectively. These weights indicate that the single predictor and criterion measures which account for the greatest portion of the variance common to the two batteries are the Metro N predictor and the Edmonton Arithmetic criterion. Consequently, the overlap of variance between the two batteries seems to be very largely a function of numerical skills. (The lesser contribution of the Metro R Test to the common variance is reflected in its relatively low canonical weight of .528).

The low weight associated with the Monroe Test indicates that it is contributing little to the common variance between the two batteries. In view of the high correlation between this test and both Teacher Ratings and Edmonton Arithmetic (.716 and .636 respectively), and its low correlation with both predictors (.312 and .389), it seems that almost all of the common variance accounted for by the Monroe Test is also common to the other criteria, which also include variance not associated with the Monroe Test. Thus, the Monroe Test makes only a slight individual contribution



to the variance common to both predictor and criterion batteries, and, in a statistical sense, its value is slight.

5. On the basis of the present procedures, two kinds of error may occur when children are selected for the first grade. They may be selected only to be decelerated, and children who would have succeeded may be rejected. In the present study, the total of both errors was about 13.5 per cent. If the relationship between predictor and criterion batteries were purely random (i.e.  $R_{\text{canonical}} = 0$ ), the number of incorrect rejections associated with the same cut-off points on both batteries would have risen from 38 to 53, or 16.8 per cent (Arbous).<sup>53</sup> Consequently, it can be seen that the selection procedures were considerably better than chance, and Hypothesis E may be said to have been supported.

It should be noticed that errors involving incorrect rejection of candidates could be reduced by lowering the cut-off score on the predictors required for admission, and would disappear altogether if all candidates were admitted. On the other hand, the other class of error (admission of students subsequently decelerated) would increase accordingly. Similarly, the proportion of decelerates could be sharply reduced by raising the cut-off score on the predictors, with a corresponding increase in the number of incorrect rejections.

Clearly, the position of the cut-off point will be a function of the value judgements made about each kind of error.

6. There were no significant correlations between test scores and age of the subjects.

---

<sup>53</sup> Arbous, op. cit., page 46.



### Conclusions

1. The value of the Metropolitan Readiness Test as a predictor of first-grade achievement in West Jasper Place Public Schools has been shown to be significantly better than chance. Prediction is substantially short of perfect and it is conceivable that some other better and more efficient method may be found in the foreseeable future. However, final assessment of the overall value of the present testing program depends upon seriousness with which errors of placement are regarded and the cost of improved techniques. The pre-school testing program for admission of under-age candidates to first grade of the West Jasper Place Public School System is therefore satisfactory so long as present proportion of misplacement is regarded as satisfactorily low.

2. Fifty-nine candidates were misplaced by the predictors (proportion of incorrect selections is 13.5 per cent). Of these, only 38 represent rejected under-age candidates that would have been successful had they been admitted to first grade.

3. What the Monroe Reading Test measures is better measured by the other criteria. An argument can therefore be advanced that the Marion Monroe Reading Test is actually of very little use in the context of the other tests, and therefore its omission would not seriously hamper the selection of students.

It is probable, and significant that Teacher Ratings measure the same achievement factors as are measured by the Monroe Test. Whether this is the case because teachers are compiling their ratings after results of the Monroe Test are known, or whether Teacher Ratings are a more valid method of measuring achievement as defined, retention of the Marion Monroe



Reading Test as the only present objective criteria of reading ability is no doubt desirable, and more politic than discarding it in favor of subjective Teacher Ratings.

4. Teacher Ratings correlate highly with Edmonton Public School Board Arithmetic Test (Grade 1) scores as an effective measure of first-grade achievement in the West Jasper Place Public School System, as might be expected as a result of "halo effect." Furthermore, there was a significant relationship between Grade One teachers' ratings of student performance and sex of the student; a relationship which was not significant in the case of the more objective measures. This suggests that grade one teachers tend to overestimate the performance of the girls in their classes; a tendency which has been reported elsewhere (page 23). The absence of significant relationship between sex and achievement on the objective measures is consistent with the findings of the American Bureau of Education (page 23), while the tendency for scores on the Monroe to reach low level of significance after one year of school is also consistent with other studies (page 23).

5. None of the correlations of test scores with age of the subjects employed in the study was significant. This suggests that the test scores are independent of chronological age. This finding is consistent with that of Bevington (page 19).

#### Recommendations

The pre-school testing program of under-age candidates for admission to first-grade in the West Jasper Place Public Schools appears to be as accurate as can be expected, and should be maintained in its present form until such time as more effective predictors of achievement become available.



## BIBLIOGRAPHY

- Ammens, M. P. and Goodlad, J. I. "When to Begin". Childhood Education. Association for Childhood Education International, Washington, 1955.
- Arbous, A. G. Tables for Aptitude Testers. Capetown: South African Council for Scientific and Industrial Research, 1952.
- Baker, Emily. "Reading Readiness is Still Important," Elementary English, Volume XXII, January, 1955.
- Broom, M. E. and others. Effective Reading Instruction. New York: McGraw-Hill Book Company Inc., 1951.
- Bevington, W. "Effect of Age of Entrance into Grade One on Subsequent Achievement." Unpublished Master's Thesis, University of Alberta, 1957.
- Carroll, Marjorie W. "Sex Differences in Reading Readiness." Master's Thesis. Boston: Boston University, 1941.
- Cone, H. R. "Brookline Admits Them Early," Nation's Schools, March, 1955.
- Dean, Charles D. "Predicting First Grade Achievement." Elementary School Journal, 1939.
- Dey, Jean D. Theory and Practice Governing the Time of School Entrance. Monographs in Education No. 4. The Alberta Advisory Committee on Educational Research. Edmonton: University of Alberta, 1960.
- Donnelly, Helen E. "A Study of Word Recognition Skills in Grade One," Master's Thesis, Boston: Boston University, 1940.
- Fast, Irene. "Kindergarten Training and Grade One Reading," Journal of Educational Psychology, Volume XLVIII, 1957.
- Fuller, Elizabeth M "Peas in a Pod." Educational Leadership, 1954.
- Gates, Arthur I. "Readiness for Beginning Reading." Readiness for Learning. Hildreth, G. (comp.). Bulletin of the Association for Childhood Education: Washington, 1941.
- Gates, Arthur I. "The Necessary Mental Age for Beginning Reading." Elementary School Journal, 1937.



- Gilmartin, Catherine E. "Progress of Under-Age Children Admitted by Test to First Grade in Quincy Schools." Master's Thesis, Boston: Boston University, 1946.
- Harrison, M. Lucille. Reading Readiness. Cambridge, Mass: Houghton, Mifflin Co., 1939.
- Hayes, Eleanor. "Why Pupils Fail." Educational Method. October, 1933.
- Hildreth, Gertrude. Readiness for School Beginners. Yonkers-on-Hudson, New York: World Book Co., 1950.
- Hobson, James R. "Mental Age as a Workable Criterion for School Admission." Elementary School Journal, 1948.
- Hurlock, E. Child Growth and Development. New York: McGraw-Hill Book Company Inc., 1949.
- King, Inez B. "Effect of Entrance Age Into Grade One Upon Achievement in Elementary School," Elementary School Journal, Volume LV, February, 1955.
- Lazar, May. Determining Readiness for Reading. Educational Research Bulletin No. 6. Board of Education of the City of New York, 1943.
- "Metropolitan Readiness Tests - Validity," New York and Tarrytown: Harcourt, Brace and World Inc., 1949.
- "Metropolitan Readiness Tests - Using the Test Results." New York and Tarrytown: Harcourt, Brace and World Inc., 1949.
- Morphett, M. V. and Washburne, C. "When Should Children Begin to Read?" Elementary School Journal, 1931.
- Olsen, Willard C. "Reading as a Function of the Total Growth of the Child." William S. Gray (comp. and ed.) Reading and Pupil Development. Supplementary Educational Monographs No. 5, Chicago: University of Chicago Press, 1940.
- "Pupils' Readiness for Reading Instruction upon Entrance to First Grade." City School Leaflet No. 23, U.S. Department of the Interior. Washington: Bureau of Education, December, 1926.
- Reed, Mary M. "An Investigation of Practice for the Admission of Children and the Promotion of Children from First Grade." Doctor's Dissertation, New York Teacher's College, Columbia University, 1927.
- Rice, W. D. "Some of the Educational Gains and Losses to Children Who Start to School at Different Ages." Master's Thesis. Chapel Hill: Duke University, 1934.



- Rosebrook, Wilda. "Preventing Reading Deficiency." Elementary School Journal, 1936.
- Samuels, F. L. "Sex Differences in Reading Achievement." Journal of Educational Research, 1936.
- Smith, Nila B. "Readiness for Reading, Parts I and II," Elementary English, Volume XXVII, 1950.
- Smith, Nila B. "Readiness for Reading." Readiness for Reading and Related Language Arts. Bulletin, National Conference on Research in English, Elementary English, Chicago, 1950.
- Traxler, A. E. The Reliability and Predictive Value of the Metropolitan Readiness Tests. Educational Records Bulletin, No. 47, New York: Educational Records Bureau, February, 1947.
- Taylor, Christian D. The Effect of Training on Reading Readiness. Studies in Reading, Volume II. Scottish Council for Research in Education. London: University of London Press, 1948.
- Witty, Paul. "What is Basic in Readiness for Learning?" Readiness for Learning, Hildreth, G. (comp.) Bulletin of the Association for Childhood Education: Washington, 1941.













**B29818**